DETECCIÓN Y BIOMARCADORES

Depression risk architecture: from etiology to algorithmic screening

SentirIA Research Papers · 2026

Infraestructura de detección temprana en salud mental

Este documento es parte de la base científica de SentirIA,
plataforma de detección temprana y monitoreo continuo de deterioro en salud mental.
No constituye diagnóstico clínico. La evaluación es responsabilidad del profesional.

Depression risk architecture: from etiology to algorithmic screening

Depression is a multifactorial condition driven by interacting biological, psychological, social, and environmental forces — and emerging evidence now supports integrating these signals into computational risk algorithms with clinically meaningful accuracy. The strongest individual risk factors include prior depressive episodes (OR = 4.4), high discrimination exposure (OR = 5.39), childhood adversity with 4+ ACEs (OR = 3.08), and social isolation via Mendelian randomization (OR = 3.70). Machine learning models combining multiple signals achieve pooled AUC of ~0.80 for depression prediction, while multi-modal systems fusing text, voice, and behavioral data reach AUC = 0.95. For a WhatsApp-based screening tool like SentirIA, conversational PHQ-9 administration shows ICC ≥ 0.91 with standard versions, and validated text and vocal biomarkers can serve as continuous passive monitoring layers between active assessments.

I. Biological substrates: genetics, inflammation, and the gut-brain axis

Depression's heritability sits at approximately 37% (95% CI: 31–42%) from twin study meta-analyses, with first-degree relatives facing a 2- to 3-fold elevated risk. The genetic architecture is highly polygenic — a landmark GWAS meta-analysis (Howard et al., 2019, *Nature Neuroscience*) identified 102 independent variants, though individual locus odds ratios remain below 1.2. Polygenic risk scores (PRS) show more practical utility: individuals in the top PRS decile carry OR = 2.02 (95% CI: 1.67–2.42) for depression, and high genetic risk combined with unfavorable lifestyle yields HR = 2.18 (95% CI: 1.84–2.58) compared to low-risk individuals with favorable lifestyles. Critically, favorable lifestyle offsets roughly half of high genetic risk (HR = 0.51), establishing lifestyle as a modifiable buffer even against strong genetic loading.

The serotonin hypothesis has undergone significant revision. Moncrieff et al.'s 2022 umbrella review in *Molecular Psychiatry* concluded there was "no consistent evidence" for lowered serotonin activity causing depression. However, Jauhar et al. (2023) and 36 co-authors published a substantial rebuttal, citing PET studies showing reduced brain serotonin release in depression. The current consensus frames depression as a dysbalanced dynamic system involving serotonin, dopamine, norepinephrine, glutamate, and GABA — not a simple chemical deficit. SSRIs nonetheless show efficacy across 21 drugs (Cipriani et al., 2018, *Lancet*), though only ~33% of patients achieve remission with SSRI monotherapy.

Neuroinflammation defines an important subtype. Meta-analyses find consistently elevated IL-6 (d = 0.54), CRP (d = 0.54), and TNF-α (d = 0.40, 95% CI: 0.15–0.65) in depressed patients (Osimo et al., 2020, *Molecular Psychiatry*). However, only about 25% of depressed patients exhibit chronic low-grade inflammation (CRP >3 mg/L), suggesting inflammation characterizes a subtype

rather than a universal mechanism. This subtype shows poor SSRI response but better outcomes with ketamine, ECT, and anti-inflammatory agents. HPA axis dysregulation compounds this picture: approximately 40–60% of depressed patients show abnormal cortisol non-suppression on dexamethasone suppression testing.

The gut-brain axis adds another biological layer. Meta-analyses consistently find reduced Coprococcus and Faecalibacterium in MDD patients. Germ-free mice show 2.8-fold reduction in serotonin levels, and probiotic effects (L. rhamnosus improving mood and decreasing corticosterone) are abolished in vagotomized mice, confirming vagal mediation. Epigenetic mechanisms — particularly BDNF promoter hypermethylation, NR3C1 methylation impairing HPA axis feedback, and FKBP5 intronic hypomethylation sustaining glucocorticoid resistance — provide pathways for intergenerational stress transmission.

II. Psychological and social risk factors with quantified effect sizes

Neuroticism stands as the single strongest personality predictor of depression. Kotov et al.'s 2010 *Psychological Bulletin* meta-analysis (175 studies) found d = 1.33 for depression, while Mendelian randomization evidence shows each additional neuroticism scale point corresponds to a 1.28-fold higher OR for depression — providing evidence of causal influence, not mere correlation. Rumination shows a large meta-analytic effect size for predicting psychopathology, exceeding avoidance, suppression, and reappraisal (Aldao et al., 2010, *Clinical Psychology Review*). The brooding subtype shows correlations of r = 0.51–0.53 with depressive worry.

Adverse Childhood Experiences demonstrate a clear dose-response gradient. From the original Felitti et al. (1998) study through recent prospective meta-analyses (Thurston et al., 2025, 62 studies from 15 countries), the evidence is consistent:

- 1 ACE: OR = 1.43 for increasing depression trajectory
- 2–3 ACEs: OR = 1.97 (95% CI: 1.30–3.00)
- 4+ ACEs: OR = 3.08 (95% CI: 1.86–5.09)

Population-attributable fraction estimates suggest preventing ACEs could reduce depression prevalence by 44.1% (Merrick et al.), making childhood adversity the single most impactful modifiable risk factor at a population level. Insecure attachment compounds this risk, with avoidant attachment showing g = 0.83 for MDD in transdiagnostic meta-analysis.

Social determinants carry substantial and often underestimated weight. Loneliness predicts new-onset depression with pooled adjusted OR = 2.33 (95% CI: 1.62–3.34). Mendelian randomization pushes the estimated causal effect of social isolation even higher, at OR = 3.70 (95% CI: 2.32–5.89). Low income shows OR = 1.96 (95% CI: 1.53–2.52) — the strongest socioeconomic association — and notably, this effect is comparable across high- and low-income countries,

indicating relative income matters more than absolute income. Adults reporting high discrimination carry OR = 5.39 (95% CI: 3.61–8.04) for positive depression screening.

III. Specific contributing factors: sleep, seasons, hormones, and chronic illness

Sleep disruption is among the most actionable risk factors. Three successive meta-analyses converge: insomnia confers roughly 2- to 3-fold risk for incident depression (Baglioni et al., 2011: OR = 2.60; Li et al., 2016: RR = 2.27; Hertenstein et al., 2023: confirmation). The relationship is bidirectional — depression predicts insomnia and vice versa — but treating insomnia with CBT-I concurrently improves both conditions. Sleep duration follows a U-shaped risk curve, with both short and long sleep increasing depression risk.

Seasonal affective disorder prevalence ranges from 1.5% to 10% globally, with a consistent latitude gradient: each 1-degree latitude increase raises SAD prevalence by 0.2 percentage points and subsyndromal SAD by 0.32 percentage points (Kim et al., 2025 meta-analysis, 24 studies, 32,866 participants). Light therapy at 2,500–10,000 lux for 30–60 minutes daily is first-line treatment. Vitamin D supplementation shows a moderate aggregate effect on depression (Hedges' g = −0.317), with stronger effects in those with MDD (g = −0.729) and a dose-response of SMD = −0.32 per 1,000 IU/day.

Chronic illness comorbidity rates are striking across conditions:

Condition	Depression prevalence	Key risk metric
Type 2 diabetes	25–35%	24% increased risk vs. general population
Post-MI / cardiovascular disease	20–30%	OR = 4.5 for MI in depressed patients
Cancer	Up to 67% in some clinical samples	Varies by cancer type
Fibromyalgia	25% current; >50% lifetime	5× more likely to have MDD
Multiple sclerosis	23.7%	Among most prevalent MS comorbidities
Systemic lupus erythematosus	24–39%	12% annual incidence of new depression

Shared biological pathways — particularly pro-inflammatory cytokines, HPA axis dysregulation, and reduced BDNF — underpin these comorbidities. Pooled across chronic physical diseases, the overall OR = 3.1 (95% CI: 1.8–5.2) for anxiety and/or depression versus controls.

Hormonal transitions create windows of vulnerability. Postpartum depression affects 10–15% of

women, driven by the dramatic postpartum crash in estrogen and progesterone (which increase up to 100-fold during pregnancy) and in allopregnanolone (which increases ~40-fold). Perimenopause independently raises depression risk 2- to 4-fold even in women with no prior history (Harvard Study of Women's Moods/SWAN data). A Danish cohort of 188,648 first-time mothers found that hormonal-contraceptive-associated depression predicted higher PPD risk, supporting the concept of a "hormone-sensitive" phenotype that spans the reproductive lifespan.

Substance use comorbidity follows predictable patterns: alcohol use disorder is present in 20.8% of MDD patients (36% of men), with Mendelian randomization evidence suggesting the comorbidity primarily reflects self-medication (depression → alcohol use). Cannabis use confers OR = 1.29 (95% CI: 1.13–1.46) for depression overall, rising to OR = 1.62 for heavy users.

IV. Contextual and environmental signals: what the calendar and clock reveal

Weather effects on mood are real but small in the general population. Denissen et al.'s multilevel study found regression weights not exceeding ~0.07 for individual weather parameters. However, at extremes the effects become clinically significant: extreme high temperatures (99th percentile) show RR = 1.18 (95% CI: 1.08–1.29) for mental health disorders, and each 1°C temperature rise links to a 1% increase in suicides. Sunlight exposure shows a clearer dose-response: UK Biobank data suggests an optimal protective threshold of ~1.5 hours daily outdoor time, with each additional hour conferring OR = 0.96 (95% CI: 0.92–0.98) for lifetime depression risk. A Dutch cohort study found each additional hour of daylight exposure reduced depression odds by 7% in late chronotypes.

Individual differences are enormous — a Nature (2025) mobile health study identified distinct subgroups showing opposite temperature-depression correlations. This heterogeneity means population-level weather signals should be personalized in any algorithm.

The holiday-suicide link is a persistent myth. CDC data consistently shows December has among the lowest daily suicide rates, while rates peak between April and August. Suicide runs 11–23% higher in spring and summer. The spring paradox — increased suicide during apparent renewal — may reflect increased energy enabling action on suicidal thoughts, circadian misalignment, and the contrast between others' improvement and continued personal suffering. Despite 25 years of debunking by the Annenberg Public Policy Center, 80% of Americans still believe the myth. Holiday stress is real (41% of adults report increased holiday stress per APA polling; 24% of those with mental illness say holidays make their condition "a lot" worse), but it does not translate to increased clinical episodes or suicide.

Weekday-weekend patterns are robust. Stone et al. (2012), analyzing 340,000 Americans, found strong support for better mood on weekends and Fridays, regardless of demographic factors. Golder & Macy's analysis of 509 million tweets from 2.4 million users across 84 countries

replicated this globally. The biggest mood decrease occurs from Sunday to Monday, consistent with "Sunday scaries" — though absolute Monday mood is no worse than Tuesday through Thursday. The weekend effect persists even in retired individuals, suggesting factors beyond employment contribute.

Diurnal mood variation is a DSM criterion for melancholic features. The STAR*D study found DMV in 22.4% of depressed patients, with 31.9% showing morning worsening. The "Mind After Midnight" hypothesis (Perlis et al., 2022) proposes that nocturnal wakefulness produces attentional biases, negative affect, altered reward processing, and prefrontal disinhibition that synergistically promote behavioral dysregulation. Golder & Macy's Twitter data shows negative feelings peak between 9 PM and 3 AM.

Anniversary reactions and birthday effects carry real risk. Japanese data (2 million death records) show people are 50% more likely to die by suicide on their birthday. Males in England and Wales show IRR = 1.39 (95% CI: 1.18–1.64) for birthday suicide. Milestone birthdays (20, 30, 40, 60) carry even higher risk. Bereaved spouses show elevated despair and depressive symptoms during the month of the deceased's birthday, though these effects attenuate after the first 6 months.

V. Algorithmic approaches to multi-signal depression risk assessment

Several validated clinical prediction models provide benchmarks. PredictD (Bellón et al., 2011), developed across 6 European primary care sites, achieves AUC = 0.79 for 12-month major depressive episode onset, using age, sex, physical/mental health, life events, social support, and prior depression. The adolescent IDEA-RS model achieves AUC = 0.72–0.73 on external validation. Machine learning meta-analyses show pooled accuracy of 0.76 and AUC = 0.80 across 155 studies (mean sensitivity = 0.73, specificity = 0.75), though study quality inversely correlates with reported performance.

The most important features in ML depression models, ranked by consistent importance across studies, are: (1) feeling loved/wanted and social connectedness, (2) prior depressive symptoms/baseline severity, (3) sleep disturbances, (4) physical activity levels, (5) stress and traumatic life events, (6) demographics (age, female sex), (7) comorbid anxiety, (8) social support, (9) chronic disease comorbidities, and (10) genetic risk scores (marginal improvement: AUC from 0.66 → 0.68).

Digital phenotyping provides continuous passive signals. Saeb et al.'s seminal 2015 study found GPS-derived circadian movement regularity correlated $r = -0.63$ with PHQ-9 scores, location entropy at $r = -0.58$, and phone usage duration at $r = 0.54$. Classification accuracy reached 86.5% using normalized entropy alone. The BiAffect study demonstrated that keystroke metadata (typing speed, autocorrect rate, backspace patterns) correlates with depression and mania

severity, with phone movement while typing predicting anhedonia (β = −0.12, p = 0.0003). The RADAR-CNS project (623 participants with recurrent MDD) confirmed cross-sectional relationships between digital phenotypes and PHQ-8 but found longitudinal relationships weaker and heterogeneous — depression manifests in distinct behavioral subgroups.

For combining heterogeneous signals, the strongest published framework is a Bayesian Network architecture (Thymia/PMC, 2025): paralinguistic and linguistic features feed surrogate neural networks producing symptom-level predictions, which are then integrated through a Bayesian Network. Performance on 30,135 speakers: depression AUC = 0.842, anxiety AUC = 0.831, with excellent calibration (ECE = 0.018). Key advantages include transparency, handling of missing data, clinician-interpretable outputs, and the ability to integrate across noisy multimodal inputs. Multi-modal fusion approaches consistently outperform unimodal: pooled AUC = 0.95 (95% CI: 0.92–0.96) versus 0.84–0.92 for single modalities.

VI. Building SentirIA: practical evidence for a WhatsApp-based screener

Conversational PHQ-9 is psychometrically sound

Multiple validation studies confirm chatbot-administered PHQ-9 performs equivalently to standard self-report. The Perla conversational agent (Spanish-language) achieved sensitivity = 96%, specificity = 90%, and Pearson correlation of r = 0.91 with standard PHQ-9. HopeBot (GPT-4o powered, voice-based) showed ICC = 0.91 with self-administered versions; 45% of scores were identical. Completion rates exceed 98%. Adaptive approaches using Item Response Theory (e.g., CAT-DI) reduce item burden to 5–7 questions while maintaining comparable reliability (r = 0.81 with PHQ-9) and sensitivity (0.92) for MDD detection.

Text and voice biomarkers provide passive monitoring layers

Linguistic markers of depression are well-established. Absolutist language ("always," "nothing," "completely") is the single strongest text marker, with Cohen's d = 3.85–4.41 for suicidal ideation forums versus controls (Al-Mosaiwi & Johnstone, 2018). First-person singular pronoun usage shows a smaller but robust effect (r = 0.13, 95% CI: 0.10–0.16). NLP meta-analysis across 123 studies yields pooled accuracy of 0.80 and AUC of 0.79 for text-based depression detection, with transformer models (BERT, MentalBERT) achieving state-of-the-art performance. A BERT-LSTM hybrid detected concerning mental health changes 6.3 weeks before users explicitly mentioned problems.

Vocal biomarkers from audio messages offer complementary signals. Systematic reviews report AUC ranging from 0.71 to 0.93 for voice-based depression classification. Key features include reduced fundamental frequency (F0), flattened pitch modulation, increased jitter/shimmer, longer and more frequent pauses, and slower speaking rate. Multi-modal methods combining audio with

text achieve pooled AUC = 0.95. WhatsApp voice messages (opus codec) require minimum 20–25 seconds of speech and conversion to 16-kHz linear PCM for reliable acoustic analysis; compression may degrade some spectral features.

Recommended signal architecture and weighting

Based on the evidence, a practical SentirIA architecture should weight signals in this hierarchy:

1. PHQ-9 conversational scores — highest weight (validated gold standard, AUC ≈ 0.84)
2. Linguistic/text features — high weight (pooled AUC = 0.79, continuous passive collection)
3. Vocal biomarkers — moderate weight (AUC = 0.71–0.93, intermittent from voice notes)
4. Behavioral metadata — contextual weight (message timing, frequency, response latency)
5. Environmental/contextual signals — supportive weight (time of day, season, calendar events)

Use late fusion or Bayesian Network integration for robustness to missing modalities. Train with modality dropout so the system degrades gracefully when voice messages are absent. Implement exponential time-weighting favoring recent signals, track individual baselines, and detect deviations from personal norms rather than population norms. Calibrate probability outputs across demographic subgroups using Platt scaling; address gender bias (3 of 4 studies found higher male misclassification) and cultural/linguistic variation (most models are English-trained; multilingual BERT variants are needed for Spanish deployment).

Regulatory and ethical landscape

No AI-enabled medical devices for mental health have received FDA authorization as of late 2025, despite 1,200+ AI devices cleared in other medical areas. The FDA's Digital Health Advisory Committee expressed comfort with screening for mild symptoms but greater concern about moderate-severe depression. Positioning SentirIA as a clinician-adjunct clinical decision support tool may qualify for enforcement discretion. Crisis escalation protocols (mandatory for PHQ-9 Item 9 suicidal ideation flags) and human oversight are regulatory prerequisites. The WHO's mhGAP Third Edition (2023) explicitly endorses digital self-help tools, and WHO's Step-by-Step program provides a model for culturally adaptable deployment in low-resource settings.

Conclusion: toward precision screening through signal integration

The evidence base supports three actionable conclusions for computational depression risk assessment. First, no single signal dominates — the strongest individual risk factors (prior episodes, childhood adversity, social isolation, neuroticism) each capture only a fraction of variance, but multi-modal integration consistently pushes AUC above 0.90. Second, temporal dynamics matter more than static snapshots — within-person change trajectories, circadian

patterns, and deviation from individual baselines are more clinically informative than population-normed cutoffs. The RADAR-CNS project's finding that depression manifests in distinct behavioral subgroups underscores the need for personalization. Third, the regulatory path favors screening over diagnosis — positioning AI tools as decision support that routes to human clinicians carries lower risk classification and leverages the strongest evidence base (conversational PHQ-9 ICC > 0.90, multi-modal AUC = 0.95). The convergence of validated chatbot screening, passive linguistic and acoustic monitoring, and digital phenotyping makes WhatsApp-based platforms a viable delivery vehicle — particularly for populations in low- and middle-income countries where traditional mental health infrastructure remains scarce.